Genetic analysis and molecular mapping of a pale flower allele at the *W4* locus in soybean

Min Xu and Reid G. Palmer

Abstract: In soybean (*Glycine max* (L.) Merr.), the *w4*-mutable line that harbors the *w4-m* allele was identified in 1983. It was proposed that this line contained an autonomous transposable element at the *W4* locus, which is a major locus controlling the biosynthesis of anthocyanin. The *w4-m* allele can revert to the *W4* allele that produces the wild-type phenotype, or sometimes to other alleles that produce intermediate phenotypes. Mutant plants that produce pale flowers were identified among the progeny of a single germinal revertant event from the *w4*-mutable line. Through genetic analysis, we established that the pale-flower mutation was conditioned by a new allele (*w4-p*) at the *W4* locus. The *w4-p* allele is dominant to the *w4* allele but recessive to the *W4* allele, and the *w1* allele has an epistatic effect on the *w4-p* allele. The pale-mutant line (*w4-pw4-p*) was designated as Genetic Type Collection number T369. An F_2 mapping population derived from the cross of Minsoy (*W4W4*) × T369 (*w4-pw4-p*) was used to map the *W4/w4-p* locus, using simple sequence repeat (SSR) markers. The *W4* locus was located at one end of molecular linkage group D2, 2.3 cM from the SSR marker Satt386 and close to the nearby telomere.

Key words: Glycine max, *w4*-mutable line, transposable element, SSR markers

Résumé : Chez le soja (*Glycine max* (L.) Merr.), la lignée "*w4*-mutable" contenant l'allèle *w4-m* a été identifiée en 1983. Cette lignée contiendrait un élément transposable autonome au locus *W4*, lequel est un des principaux locus qui contrôlent la synthèse de l'anthocyane. L'allèle *w4-m* peut muter spontanément pour restaurer l'allèle *W4*, lequel confère un phénotype sauvage, ou encore produire d'autres allèles de phénotype intermédiaire. Des mutants à fleurs pâles ont été identifiés parmi la progéniture d'un révertant germinal issu de la lignée "*w4*-mutable". Par analyse génétique, les auteurs ont montré que les fleurs pâles sont dues à un nouvel allèle (*w4-p*) au locus *W4*. L'allèle *w4-p* est dominant par rapport à *w4* mais récessif par rapport à *W4*. De plus, l'allèle *w1* a un effet épistatique sur l'allèle *w4-p*. La lignée à fleurs pâles (*w4-pw4-p*) s'est vue attribuer le numéro T369 dans la 'Genetic Type Collection'. Une population de cartographie F_2 issue d'un croisement entre Minsoy (*W4W4*) et T369 (*w4-pw4-p*) a été employée pour localiser le locus *W4/w4-p* à l'aide de marqueurs microsatellites (SSR). Le locus *W4* est situé à une extrémité du groupe de liaison D2, à 2,3 cM du marqueur Satt386 et à proximité du télomère.

Mots clés : Glycine max, lignée "*w4*-mutable", élément transposable, marqueurs microsatellites.

[Traduit par la Rédaction]

Introduction

In soybean, 6 genes, *W1*, *W2*, *W3*, *W4*, *Wp*, and *Wm*, have been identified that condition anthocyanin pigmentation in flowers and hypocotyls (Hartwig and Hinson 1962; Stephens and Nickell 1992; Palmer et al. 2004). The *W4* locus has a major effect on the biosynthesis of anthocyanin in soybean flowers and hypocotyls. Mutations at the *W4* locus lead to less anthocyanin pigmentation in both tissues. In the *W1*_ genetic background, soybean lines with genotype *W4*_ generate wild-type purple flowers and purple hypocotyls. Soybean lines with genotype *w4w4* (a stable recessive mutation at the *W4* locus) generate near-white flowers and green hypocotyls (Groose and Palmer 1991).

A mutable allele (*w4-m*) that conditions unstable anthocyanin pigmentation in soybean flowers and hypocotyls was identified at the *W4* locus. The genetic stock for the inbred mutant line (*w4-mw4-m*) was registered as the *w4*-mutable line, and was assigned Genetic Type Collection number T322 (Palmer et al. 1990). It has been proposed that the *w4*-mutable line contains an autonomous transposable element at or near the *W4* locus, which results in the *w4-m* allele (Palmer et al. 1989). The *w4-m* allele can revert to the *W4* allele after excision of the putative transposable element. Somatic reversion of the *w4-m* allele results in plants that produce variegated flowers and hypocotyls (Groose et al. 1988). Germinal reversion of the *w4-m* allele, in most cases, results in plants that produce all wild-type purple flowers and hypocotyls (Groose et al. 1990). Many new mutations were identified among self-pollinated progenies of these wild-type germinal revertants, for example, chlorophyll-deficient

Received 30 March 2004. Accepted 27 September 2004. Published on the NRC Research Press Web site at http://genome.nrc.ca on 18 February 2005.

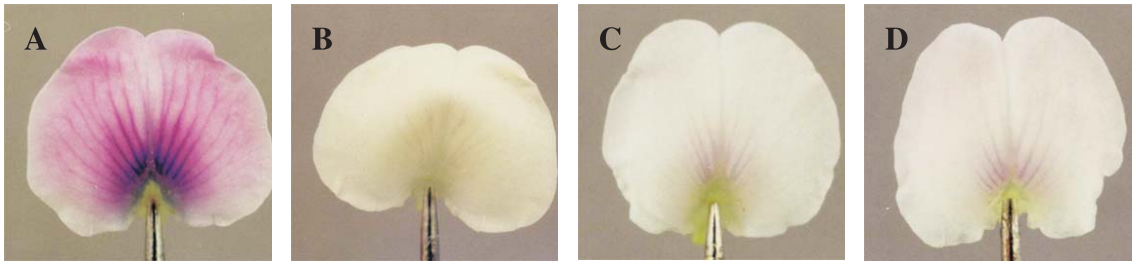
Corresponding Editor: G.J. Scoles.

M. Xu. Department of Agronomy and Interdepartmental Genetics Graduate Program, Iowa State University, Ames, IA 50011, USA.
R.G. Palmer.[1] USDA–ARS–CICGR and the Department of Agronomy and Interdepartmental Genetics Graduate Program, Iowa State University, Ames, IA 50011, USA.

[1]Corresponding author (e-mail rpalmer@iastate.edu).

 doi: 10.1139/G04-105

Fig. 1. Standard petals of flowers of soybean wild-type plants and mutants. (A) Wild-type purple Harosoy (*W1W1W4W4*); (B) near-white Harosoy *w4* isoline (L72–1138) (*W1W1w4w4*); (C and D) pale-mutant line, T369 (*W1W1w4-pw4-p*).



leaves, necrotic roots, and complete or partially sterile plants (Palmer et al. 1989). Occasionally, germinal revertants may produce intermediate-colored flowers (Groose et al. 1990).

Two mutant lines with intermediate-colored flowers (dilute purple and pale) have been identified from the *w4*-mutable lines. The dilute-purple-mutant line was conditioned by a new allele at the *W4* locus (*w4-dp*), and was assigned Genetic Type Collection number T321 (Palmer and Groose 1993).

No genetic analysis of the pale-mutant line has been reported. The pale-mutant line produces pale flowers, which are lighter than wild-type purple but darker than recessive near-white ones (Fig. 1), and green hypocotyls. Genetic analysis and molecular mapping of the pale mutant should provide additional evidence to support the hypothesis that there is an active transposable element within the *w4-m* allele. If it is an allele at the *W4* locus, the pale-mutant line will provide new material that can be used to study gene expression and regulation of the *W4* gene.

In the soybean genome, simple sequence repeat (SSR) markers are highly polymorphic (Akkaya et al. 1992), randomly distributed on chromosomes (Akkaya et al. 1995), and suitable for mapping new genes because 1 pair of primers usually amplifies only 1 product from an inbred genotype with PCR (Cregan et al. 1994). An integrated genetic linkage map with 20 molecular linkage groups, including 606 SSR, 689 RFLP (restriction fragment length polymorphism), 79 RAPD (random amplified polymorphic DNA), 11 AFLP (amplified fragment length polymorphism), 10 isozyme, and 26 classical loci, has been constructed (Cregan et al. 1999). Of the 6 genes that control anthocyanin biosynthesis in soybean flowers, only 2 loci have been positioned on this map. The *W1* locus was located on molecular linkage group (MLG) *F* (Cregan et al. 1999), and the *wp* locus was located on MLG D1b+w (Hegstad et al. 2000).

Our objectives were to study the genetic relationship between the new pale mutation and the *W4* locus, and to position the *W4* locus on a soybean chromosome with SSR markers.

Materials and methods

Genetic experiments

The pale-mutant line was determined to be true breeding. It originated as a revertant from the original *w4*-mutable line (Asgrow XP2214) (Palmer et al. 1997). Based on the results in this paper, the Soybean Genetics Committee assigned gene symbol *w4-p* to the pale allele, and Genetic Type Collection number T369 to the pale-mutant line.

Minsoy (PI 27890), Harosoy *w4* isoline (L72–1138), and Harosoy *w1* isoline (L62–906) were each crossed with T369. The phenotypes and genotypes of these lines are described in Table 1. This was done to determine how many genes controlled the pale mutation, and to identify the relationship between the pale mutation and the *w4* locus and between the pale mutation and the *w1* locus. Progenies from all 3 crosses were analyzed in the F_1, F_2, and F_3 generations.

Molecular mapping experiment

Plant materials

The following true-breeding soybean lines were used in this experiment: Harosoy (purple flowers), Harosoy *w4* isoline (near-white flowers), Minsoy (purple flowers), and T369 (pale flowers) (Table 1).

The mapping population, consisting of 114 F_2 individuals descended from a single F_1 seed of Minsoy (purple flowers) × T369 (pale flowers), was grown at the Bruner Farm, near Ames, Iowa, in the summer of 2002. Plants were sampled, identified, and threshed individually. Twenty-four $F_{2:3}$ descendants from each F_2 plant were planted for the progeny test in the fall of 2002, in the USDA greenhouse on the Iowa State University campus, to determine the genotype of each F_2 plant.

DNA extraction and PCR conditions

DNA was extracted from freeze-dried young leaves of the Harosoy *w4* isoline, Minsoy, T369, and 114 F_2 plants of the mapping population, using the CTAB (cetyltrimethylammonium bromide) method (Keim et al. 1988).

PCRs were performed in 30 µL of reaction mixture, containing 1× PCR buffer, 1.75 mmol $MgCl_2$/L, 150 µmol dNTP/L, 0.15 µmol/L primers, 50 ng genomic DNA, and 3 U *Taq* DNA polymerase (Promega, Madison, Wis.). The PCR condition was 94 °C for 45 s, 47 °C for 45 s, and 68 °C for 45 s, for a total of 32 cycles. The PCR products were evaluated by electrophoresis on 2% Agarose 3:1 (AMRESCO, Solon, Ohio) gels in 1× TBE (Tris–borate–EDTA) buffer, or on 8% polyacrylamide denaturing gels in 1× TAE (Tris–acetate–EDTA) buffer.

Denaturing PAGE analysis

PCRs were stopped with 15 µL stop solution (90% formamide, 20 mmol EDTA/L, 0.01% bromphenol blue, and 10% ficoll), and the PCR products were denatured by heat-

Table 1. Genotypes and phenotypes of soybean lines used in the genetic and molecular mapping experiments.

Lines	Genotype	Description	Flower color
Harosoy	*W1W1W4W4*	True breeding wild-type soybean	Purple
Harosoy *w4* (L72–1138)	*W1W1w4w4*	True breeding stable recessive *w4* mutant	Near white
Harosoy *w1* (L62–906)	*w1w1W4W4*	True breeding stable recessive *w1* mutant	White
T369	*W1W1w4-pw4-p*	True breeding stable revertant from the *w4*-mutable line	Pale
Minsoy (PI 27890)	*W1W1W4W4*	True breeding wild-type soybean	Purple

Table 2. Progeny tests of crosses of Minsoy (purple flowers) and Harosoy *w4* (L72–1138) (near-white flowers) with T369 (pale flowers) in the F_1, F_2, and F_3 generations.

		Segregation in the F_2 generation					Segregation in the F_3 generation						
		No. of plants					No. of $F_{2:3}$ families						
Cross*	F_1 phenotype	Pu	pa	wh	$\chi^2_{(3:1)}$	*P*	Pu	Pu/pa	pa	pa/wh	wh	$\chi^2_{(1:2:1)}$	*P*
1	All purple	181	69		0.9	0.34	66	115	69			1.67	0.43
2	All pale		1153	397	0.31	0.58			42	74	34	0.88	0.64

* Cross 1: Minsoy (purple flowers) × T369 (pale flowers); Cross 2: Harosoy *w4* (near-white flowers) × T369 (pale flowers).

Note: Pu, plants or families that produced all purple flowers; pa, plants or families that produced all pale flowers; wh, plants or families that produced all near-white flowers; Pu/pa, families consisting of purple-flower plants and pale-flower plants; pa/wh, families consisting of pale-flower plants and near-white-flower plants.

ing at 95 °C for 2 to 4 min and then immediately incubating on ice for 5 min.

Denatured PCR products were run on 8% polyacrylamide gels (1.5 mm thickness × 28 cm length) at 60 watts for 5 h, after the gel was prerun at 70 watts for 45 min. Gel solution was prepared with 8% acrylamide/bis-acrylamide (29/1), 30% formamide, 1× TAE, and 5.6 mol urea/L. TEMED (1%) and ammonium persulphate (0.08%) were added to the gel solution just before it was poured. After electrophoresis, the gel was transferred onto a UV transparent Plexiglass plate, stained with 2× SYBR-gold stain solution (Molecular Probes, Eugene, Ore.) in the dark for at least 30 min, and observed and imaged under UV light.

Gene mapping strategies

To identify the molecular markers that are linked to the pale mutation, we used the near-isogenic lines (NILs) screening strategy (Young et al. 1988), and the bulked segregant analysis (BSA) strategy (Michelmore et al. 1991).

In the NILs screening experiment, the Harosoy and Harosoy *w4* isoline samples were screened for polymorphisms with SSR markers. In the BSA experiment, 2 bulks were constructed from the F_2 mapping population (Minsoy (purple flowers) × T369 (pale flowers)). Bulk 1 contained DNA from 10 F_2 plants homozygous for wild-type purple flowers; Bulk 2 contained DNA from 10 F_2 plants homozygous for recessive pale flowers. Homozygotes were identified by progeny tests in the $F_{2:3}$ generation.

Linkage analysis

The Linkage-1 program (Suiter et al. 1983) was used to evaluate the linkage between each candidate SSR marker and the locus that conditions the pale mutation by computing the recombination value between the pair.

The final map was constructed with the Mapmaker 2.0 program (Lander et al. 1987). Two markers were considered to be linked if the LOD score between them was equal to or higher than 3.0 and the recombination value between them was equal to or lower than 0.4. The genetic distance between 2 markers was generated from recombination rates using the Kosambi map function (Kosambi 1944).

Results

Genetic analyses

The results of progeny tests of crosses of Minsoy, Harosoy *w1*, and Harosoy *w4* with T369 in the F_1, F_2, and F_3 generations are shown in Tables 2 and 3.

F_1 plants from the cross of Minsoy (purple flowers) with T369 (pale flowers) were all wild-type purple-flower plants (Table 2, Cross 1). In the F_2 generation, the segregation ratio was 181 purple to 69 pale, which fit a 3:1 F_2 phenotypic segregation ratio ($\chi^2 = 0.90$; $P = 0.34$) (Table 2, Cross 1). In the $F_{2:3}$ families, 69 F_2 plants with pale flowers gave all pale-flower F_3 plants, 115 F_2 plants with purple flowers gave F_3 plants with purple flowers and with pale flowers, and 66 F_2 plants with purple flowers gave all purple-flower F_3 plants, which showed a 1:2:1 F_2 genotypic segregation ratio ($\chi^2 = 1.67$; $P = 0.43$) (Table 2, Cross 1). These results showed that the pale mutation was conditioned by a single locus, and that it was recessive to the wild-type purple flower phenotype.

The cross of Harosoy *w4* (near-white flowers) with T369 (pale flowers) was made to determine the relationship between the *w4* locus and the locus that conditions the pale-flower phenotype (Table 2, Cross 2). All the F_1 plants from the cross were pale flowers. In the F_2 generation, 1153 plants were pale flowers and 397 plants were near-white flowers, which showed a 3:1 phenotypic segregation ratio ($\chi^2 = 0.31$; $P = 0.58$). Among these 1550 F_2 plants, 150 plants were randomly selected to advance to the $F_{2:3}$ generation. As a result, 42 $F_{2:3}$ families were true breeding for pale flowers, 74 $F_{2:3}$ families were segregated for pale flowers

Table 3. Progeny tests of the cross Harosoy *w1* (L62–906) (white flowers) with T369 (pale flowers) in the F_2 and $F_{2:3}$ generations.

	Segregation in the F_2 generation			Family segregation in the $F_{2:3}$ generation						
	Pu	pa	wh	Pu	Pu/pa	Pu/wh	Pu/pa/wh	pa	pa/wh	wh
No. plants/families	812	245	354	11	17	19	52	9	18	39
Genotype of plants/families	*W1-W4-*	*W1*-w4-pw4-p	*w1w1*	*W1W1*	*W1W1*	*W1w1*	*W1w1*	*W1W1*	*W1w1*	*w1w1*
			—	*W4W4*	*W4w4-p*	*W4W4*	*W4w4-p*	*w4-pw4-p*	*w4-pw4-p*	—
Expected no.	794	265	353	10	21	21	41	10	21	41
	$\chi^2_{(9:3:4)}$ = 1.87	$\chi^2_{(1:2:2:4:1:2:4)}$ = 4.24								
	P = 0.39	P = 0.64								

and near-white flowers, and 34 $F_{2:3}$ families were true breeding for near-white flowers, which followed a 1:2:1 F_2 genotypic segregation ratio (χ^2 = 0.88; P = 0.64). The results (Table 2, Cross 2) suggested that the pale-flower mutation generated a new recessive allele (*w4-p*) at the *W4* locus, and that it was dominant to the previously identified recessive *w4* allele.

The results of progeny tests for Harosoy *w1* (white flowers) × T369 (pale flowers) are shown in Table 3. The F_1 plants were all purple-flower plants. The F_2 segregation was 812 purple-flower plants:245 pale-flower plants:354 white-flower plants, and showed a 9:3:4 phenotypic segregation ratio (χ^2 = 1.87; P = 0.39) (Table 3). The 165 $F_{2:3}$ families were classified into 7 categories: 1) in families producing all purple flowers, F_2 plants were genotype *W1W1W4W4*; 2) in families producing purple and pale flowers, F_2 plants were genotype *W1W1W4w4-p*; 3) in families producing purple and white flowers, F_2 plants were genotype *W1w1W4W4*; 4) in families producing purple, white, and pale flowers, F_2 plants were genotype *W1w1W4w4-p*; 5) in families producing all pale flowers, F_2 plants were genotype *W1W1w4-pw4-p*; 6) in families producing pale and white flowers, F_2 plants were genotype *W1w1w4-pw4-p*; and 7) in families producing all white flowers, F_2 plants were genotype *w1w1*_ _. The ratio followed a 1:2:2:4:1:2:4 pattern (χ^2 = 4.24; P = 0.64) (Table 3). The results showed that 2 unlinked loci, *w4-p* and *w1*, determined the flower color of the plants in this cross, and that the *w1* locus had an epistatic effect on the *w4-p* mutation. The plants with *w1w1* genotype produced white flowers, irrespective of the genotype at the *w4-p* locus.

Molecular mapping of the *W4* gene

An F_2 population with 114 individuals derived from a single F_1 plant from a cross between Minsoy (purple flowers, *W1W1W4W4*) and T369 (pale flowers, *W1W1w4-pw4-p*) was used to map the *W4* locus. The genotype of each F_2 plant was determined through a progeny test in the $F_{2:3}$ generation. The genotypic ratio of F_2 plants was 30 *W4W4*:51 *W4w4-p*:33 *w4-pw4-p*, and fit a 1:2:1 ratio (χ^2 = 1.42; P = 0.49) (Table 4).

A total of 178 SSR markers, representing all 20 MLGs of the soybean genome (Cregan et al. 1999), were selected to identify markers that were linked with the *W4* locus. The average genetic distances between any 2 adjacent markers were approximately 20 cM.

Two NILs, Harosoy (*W1W1W4W4*) and Harosoy *w4* (*W1W1w4w4*), were screened with the 178 SSR markers.

Table 4. SSR marker segregation patterns in the mapping population of Minsoy (purple flowers) × T369 (pale flowers).

	Number of F_2 plants					
Markers	AA	AB	BB	df	$\chi^2_{(1:2:1)}$	P
W4	30	51	33	2	1.42	0.49
Satt386	30	53	31	2	0.58	0.75
Satt413	31	57	26	2	0.44	0.8
Satt186	31	59	24	2	1.00	0.61
Satt574	32	60	22	2	2.07	0.36
Satt226	32	60	22	2	2.07	0.36
Satt389	36	57	21	2	3.95	0.14
Expected no.	28.5	57	28.5			

Note: AA, homozygous Minsoy genotype; AB, heterozygous genotype; BB, homozygous T369 genotype.

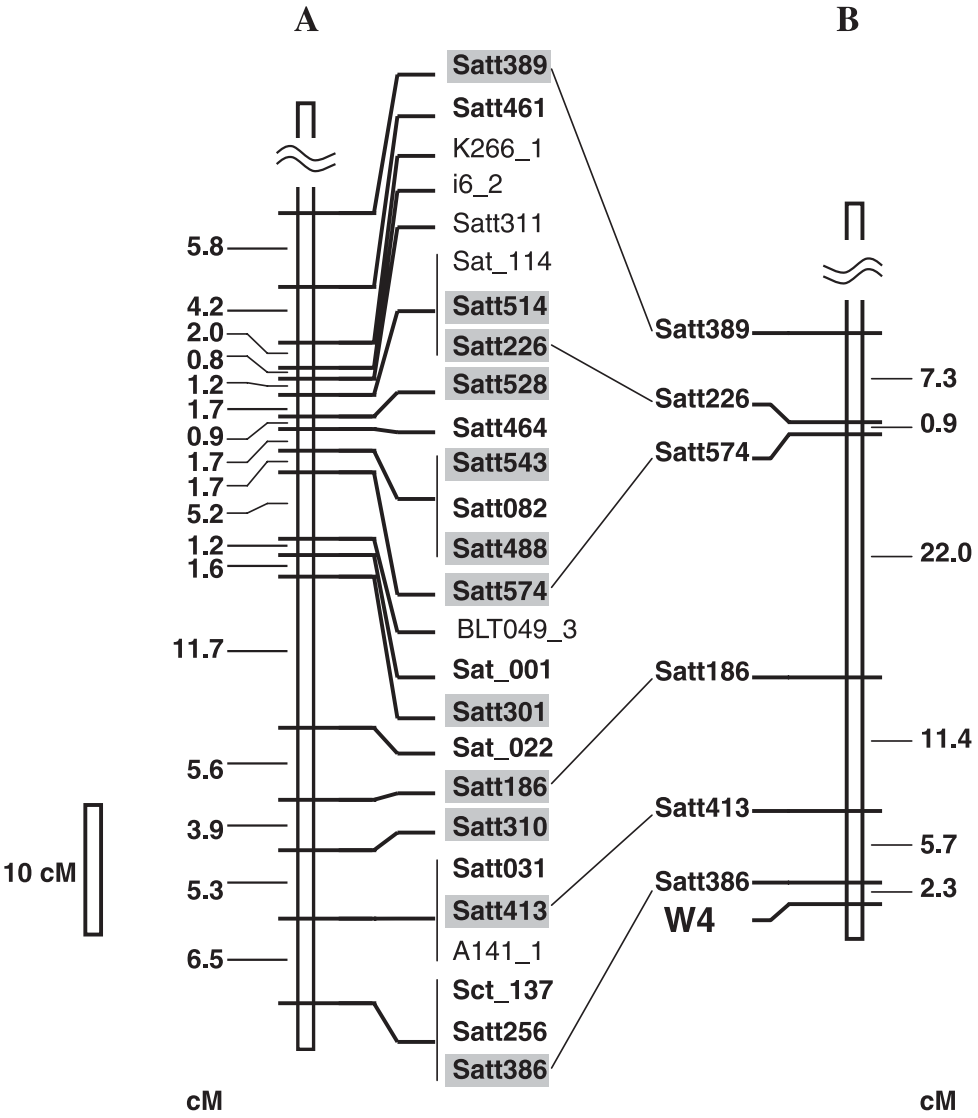
Seven SSR markers, Satt430 on MLG B1, Sat_130 on MLG C2, Satt548 on MLG D1a+Q, Satt372, Satt386, and Satt458 on MLG D2, and Satt288 on MLG G, were able to detect polymorphism between these 2 NILs on an 8% PAGE gel.

These 7 SSR markers and their flanking SSR markers were also used in the BSA experiment to screen the 2 bulks, Bulk 1 and Bulk 2, made from the F_2 mapping population, along with the 2 parental lines, Minsoy and T369. Bulk 1 contained 10 individuals homozygous for purple flowers, and Bulk 2 contained 10 individuals homozygous for pale flowers. As a result, Satt386 on MLG D2 was polymorphic between the 2 bulks and the 2 parents. The band amplified from Bulk 1 (purple flowers) was identical to the one from Minsoy (purple flowers), and the band amplified from Bulk 2 (pale flowers) was identical to the one from T369 (pale flowers). Because in BSA strategy polymorphic molecular markers between 2 bulks should represent a DNA sequence within or adjacent to the selected gene (Michelmore et al. 1991), the SSR marker Satt386 was possibly closely linked to the *W4* locus.

To confirm this, Satt386 was used to screen the 114 F_2 plants of the mapping population. The segregation ratio of Satt386 was 30 AA (Minsoy pattern):51 AB (heterozygote pattern):33 BB (T369 pattern), which fit a ratio of 1:2:1 (χ^2 = 0.58; P = 0.75) (Table 4). The recombination value between Satt386 and the *W4* locus was calculated as 0.036 ± 0.013 by the Linkage-1 program (Suiter, et al. 1983), which showed that the *W4* locus was closely linked to the SSR marker Satt386 on soybean MLG D2.

To construct a high-resolution map for the *W4* locus, an

Fig. 2. Comparison of the molecular linkage map of MLG D2 (USDA/ Iowa State University) and the molecular linkage map of the *W4* locus. Distances are shown in centiMorgans (cM). (A) MLG D2 from USDA/Iowa State University map (Cregan et al. 1999). Markers in bold are those screened for the 2 parental lines of the mapping population, Minsoy and T369. Markers highlighted in grey are the ones that detected polymorphisms between Minsoy and T369. (B) Molecular linkage map of the *W4* locus constructed in this study.



additional 18 SSR markers from MLG D2 that were at a genetic distance of less than 65 cM from Satt386 were tested for polymorphism between the 2 parental lines of the mapping population, Minsoy and T369 (Fig. 2). Eleven markers were polymorphic between these 2 parental lines (Fig. 2). Five markers, Satt389, Satt226, Satt574, Satt186, and Satt413, which could clearly identify the polymorphism between the 2 parental lines on a 2% Agarose gel, were used to screen the F_2 mapping population. The segregation ratio of each marker followed the ratio of 1 AA (Minsoy pattern) : 2 AB (heterozygote pattern) : 1 BB (T369 pattern) (Table 4).

Mapmaker 2.0 was used to generate a map for the *W4* locus, using the data from these 5 markers, along with the data from the *W4* locus and Satt386. The *W4* locus was mapped between Satt386 and the nearby telomere, with the distance of 2.3 cM between *W4* and Satt386 (Fig. 2).

Discussion

The probability of recovering new mutants is maximized by searching among progenies of germinal revertant (wild-type) plants descended from mutable plants. If the reversion of the unstable allele was the result of an excision of the element from its locus, new mutations might be detected among the progenies of these revertants. Such mutants would be expected if the excised element was inserted at a second locus.

The pale mutation was discovered among germinal revertant plants of the *w4*-mutable line. Genetic studies estab-

lished that it was controlled by a new allele (*w4-p*) at the *W4* locus. The *w4-p* allele is dominant to the *w4* allele, but recessive to the *W4* allele. The *w1* locus has an epistatic effect on the *w4-p* allele. The mutants with the *w1w1* genotype produce white flowers regardless of the genotype at the *W4* locus.

Two mutations for flowers with intermediate color were found among self-pollinated progenies of the *w4*-mutable lines. They are dilute purple (T321, *w4-dp*) (Palmer and Groose 1993) and pale (T369, *w4-p*). Both of them were identified as alleles at the *W4* locus. The molecular mechanisms are not known for these 2 alleles that evolved from the *w4-m* allele. Studies show that genes with a transposon insertion are hotspots for secondary mutations due to imprecise excision of the transposon, such as a few base changes in host sequence, and host sequence duplication or deletion. The transposon also may jump into nearby regions and transpose back to the target gene to generate new types of mutant alleles for the target gene (Walbot 2000). Thus, it is possible that the *w4-p* and *w4-dp* alleles were generated through an imprecise excision of the transposable element harbored in the *w4-m* allele, or generated by a re-insertion of the transposon that jumped from the nearby *w4-m* allele. Molecular studies are needed to address this question.

In the molecular mapping study, the *W4* locus was mapped to the terminus of MLG D2 between Satt386 and the nearby telomere, with the distance of 2.3 cM between *W4* and Satt386 (Fig. 2). Compared with the USDA/Iowa State University molecular map (Cregan et al. 1999), SSR markers, including Satt389, Satt226, Satt574, Satt186, Satt413, and Satt386, followed the same order in our map, but the genetic distances between 2 adjacent markers were shorter most of the time than the genetic distances between the same 2 markers in the USDA/Iowa State University map (Fig. 2). One reason for this may be that the parental lines of the mapping population we used were different from the parental lines used to construct the USDA/Iowa State University map. Another reason may be that the same chromosomal region is more saturated with markers in the USDA/Iowa State University map. Therefore, in the same chromosome region, the undetectable even-numbered crossovers (mainly double crossovers) in our map would be recovered in the USDA/Iowa State University map, which would lengthen their map.

The results we obtained from the molecular mapping of the *W4* locus are in agreement with the hypothesis that an active transposable element resides in the *w4*-mutable line. The *W4* locus was mapped to MLG D2; however, mutants found in independent germinal revertants of the *w4*-mutable line were mapped at different positions on different chromosomes. For example, mutant female partial-sterile 1 (*Fsp2*), female partial-sterile 2 (*Fsp3*), female partial-sterile 3 (*Fsp4*), and female partial-sterile 4 (*Fsp5*) were located on MLG C2, A2, F, and G, respectively (Kato and Palmer 2004). Mutant male-sterile and female-sterile (*st8*) was positioned on MLG J (Kato and Palmer 2003). Mutant chlorophyll-deficient leaves (CD-5) was located on MLG E (K.K. Kato and R.G. Palmer, unpublished). And, mutant chlorophyll-deficient leaves (*y20*) and mutant malate dehydrogenase1 null (*mdh1-n*) were closely linked on MLG H (unpublished data). It would be difficult to explain how the *w4-m* allele affects so many different genes *in trans* if there was no active transposable element in the *w4*-mutable line.

Transposon mutagenesis using active transposable elements is a very useful tool in gene cloning and functional genomics research in plants (Maes et al. 1999; Walbot 2000; Ramachandran and Sundaresan 2001); however, no active transposable element has ever been found in soybean. Therefore, identifying and cloning of an active transposon in soybean is very important. A good way to do this is to try to use the mutable alleles found in soybean. The mutable alleles are the genes that most likely harbor an autonomous transposable element.

Besides the *w4-m* allele, there are 3 other mutable alleles reported in soybean that may contain an autonomous transposable element: the *Y18-m* allele (Peterson and Weber 1969), the *r-m* allele (Chandlee and Vodkin 1989*a*), and the *wp-m* allele (Johnson et al. 1998). The mutable line with the *Y18-m* allele displays variegated green/yellow leaves, the mutable line with the *r-m* allele produces seeds with variegated seed coat, and the mutable line with the *wp-m* allele produces variegated flower color similar to the *w4*-mutable line.

If these 4 mutable lines do harbor an autonomous transposable element, the *w4*-mutable line and the *wp*-mutable line have more advantages than the other 2 in transposon tagging experiments. First, their germinal revertant plants are more vigorous and healthier than most of the revertants of the *Y18-m* mutable line. The *Y18-m* allele usually generates the recessive *y18* allele that produces lethal-yellow plants (Chandlee and Vodkin 1989*b*). Moreover, some of the transposition-induced new mutations, such as chlorophyll deficiency, may not be readily identified in the revertants among the *Y18-m* mutable line. Second, flower color is less affected by environments than seed-coat color; thus, the germinal revertants of the *w4-m* and *wp-m* mutable lines are easier to identify correctly than those of the *r-m* mutable line.

Finally, to clone the *w4-m* gene and the putative transposable element in the *w4-m* gene through chromosome walking, a higher-resolution map and a larger mapping population are needed. In conclusion, molecular mapping of the *W4* locus is the first step for positioned cloning of this gene, and the data provided more evidence to support the hypothesis that there is an autonomous transposable element in the *w4-m* allele.

Acknowledgements

The authors would like to thank Dr. R.W. Groose, Department of Plant, Soil and Insect Sciences, University of Wyoming, Laramie, Wyo., for providing flower pictures. The authors would also like to thank Dr. P.B. Cregan, USDA-ARS, Soybean and Alfalfa Research Laboratory, Beltsville, Md., for providing the protocols for PCR and PAGE in the SSR experiments. This is a joint contribution from the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 3769, and from the USDA, Agricultural Research Service, Corn Insects and Crop Genetics

Research Unit, and supported by the Hatch Act and the State of Iowa. The mention of a trademark or proprietary product does not constitute a guarantee or warranty of the product by Iowa State University or the USDA, and the use of the name by Iowa State University or the USDA implies no approval of the product to the exclusion of others that may also be suitable.

References

Akkaya, M.S., Bhagwat, A.A., and Cregan, P.B. 1992. Length polymorphism of simple sequence repeat DNA in soybean. Genetics, **132**: 1131–1139.

Akkaya, M.S., Shoemaker, R.C., Specht, J.E., Bhagwat, A.A., and Cregan, P.B. 1995. Integration of simple sequence repeat DNA markers into a soybean linkage map. Crop Sci. **35**: 1439–1445.

Chandlee, J.M., and Vodkin, L.O. 1989*a*. Unstable expression of a soybean gene during seed coat development. Theor. Appl. Genet. **77**: 587–594.

Chandlee, J.M., and Vodkin, L.O. 1989*b*. Unstable genes affecting chloroplast development in soybean. Dev. Genet. **10**: 532–541.

Cregan, P.B., Bhagwat, A.A., Akkaya, M.S., and Jiang, R. 1994. Microsatellite fingerprinting and mapping of soybean. Methods Mol. Cell. Biol. **5**: 49–61.

Cregan, P.B., Jarvik, T., Bush, A.L., Shoemaker, R.C., Lark, K.G., Kahler, A.L., Kaya, N., VanToai, T.T., Lohnes, D.G., Chung, J., and Specht, J.E. 1999. An integrated genetic linkage map of the soybean genome. Crop Sci. **39**: 1464–1496.

Groose, R.W., and Palmer, R.G. 1991. Gene action governing anthocyanin pigmentation in soybean. J. Hered. **82**: 498–501.

Groose, R.W., Weigelt, H.D., and Palmer, R.G. 1988. Somatic analysis of an unstable mutation for anthocyanin pigmentation in soybean. J. Hered. **79**: 263–267.

Groose, R.W., Schulte, S.M., and Palmer, R.G. 1990. Germinal reversion of an unstable mutation for anthocyanin pigmentation in soybean. Theor. Appl. Genet. **79**: 161–167.

Hartwig, E.E., and Hinson, K. 1962. Inheritance of flower color in soybeans. Crop Sci. **2**: 152–153.

Hegstad, J.M., Tarter, J.A., Vodkin, L.O., and Nickell, C.D. 2000. Positioning the *wp* flower color locus on the soybean genome map. Crop Sci. **40**: 534–537.

Johnson, E.O.C., Stephens, P.A., Fasoula, D.A., Nickell, C.D., and Vodkin, L.O. 1998. Instability of a novel multicolored flower trait in inbred and outcrossed soybean lines. J. Hered. **89**: 508–515.

Kato, K.K., and Palmer, R.G. 2003. Molecular mapping of a male-sterile, female-sterile mutant gene (*st8*) in soybean. J. Hered. **94**: 425–428.

Kato, K.K., and Palmer, R.G. 2004. Molecular mapping of four ovule lethal mutants in soybean. Theor. Appl. Genet. **108**: 577–585.

Keim, P., Olson, T., and Shoemaker, R.C. 1988. A rapid protocol for isolating soybean DNA. Soybean Genet. Newsl. **15**: 150–152.

Kosambi, D.D. 1944. The estimation of map distance from recombination values. Ann. Eugen. **12**: 172–175.

Lander, E.S., Green, P., Abrahamson, J., Barlow, A., Daly, M.J., Lincoln, S.E., and Newburg, L. 1987. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics, **1**: 174–181

Maes, T., De Keukeleire, P., and Gerats, T. 1999. Plant tagnology. Trends Plant Sci. **4**: 90–96.

Michelmore, R.W., Paran, I., and Kesseli, R.V. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. U.S.A. **88**: 9828–9832.

Palmer, R.G., and Groose, R.W. 1993. A new allele at the *w4* locus derived from the *w4-m* mutable allele in soybean. J. Hered. **84**: 297–300.

Palmer, R.G., Hedges, B.R., Benavente, R.S., and Groose, R.W. 1989. *w4*-mutable line in soybean. Dev. Genet. **10**: 542–551.

Palmer, R.G., Groose, R.W., Weigelt, H.D., and Miller, J.E. 1990. Registration of a genetic stock (*w4-mw4-m*) for unstable anthocyanin pigmentation in soybean. Crop Sci. **30**: 1376–1377.

Palmer, R.G., Harper, T.K., and Martin, S.K. 1997. A mutable flower phenotype from Asgrow XP2214. Soybean Genet. Newsl. **24**: 176–178.

Palmer, R.G., Pfeiffer, T.W., Buss, G.R., and Kilen, T.C. 2004. Qualitative genetics. *In* Soybeans: Improvement, production, and uses. 3rd ed. *Edited by* J.E. Specht and H.R. Boerma. Am. Soc. Agron. Madison, WI. pp. 137–234.

Peterson, P.A., and Weber, C.R. 1969. An unstable locus in soybeans. Theor. Appl. Genet. **39**: 156–162.

Ramachandran, S., and Sundaresan, V. 2001. Transposons as tools for functional genomics. Plant Physiol. Biochem. **39**: 243–252.

Stephens, P.A., and Nickell, C.D. 1992. Inheritance of pink flower color in soybean. Crop Sci. **32**: 1131–1132.

Suiter, K.A., Wendel, J.F., and Case, J.S. 1983. Linkage-1: a PASCAL computer program for the detection and analysis of genetic linkage. J. Hered. **74**: 203–204.

Walbot, V. 2000. Saturation mutagenesis using maize transposons. Curr. Opin. Plant Biol. **3**: 103–107.

Young, N.D., Zamir, D., Ganal, M.W., and Tanksley, S.D. 1988.Use of isogenic lines and simultaneous probing to identify DNA markers tightly linked to the *Tm-2a* gene in tomato. Genetics, **120**: 579–585.